LCA Methodology

Spy Plots

A Method for Visualising the Structure of LCA Data Bases*

[1]Ruben Huele, [2]Nico van den Berg

[1]Centre of Environmental Science (CML) Leiden University, PO Box 9518, NL-2300 RA Leiden, The Netherlands
[2]PI!MC, Schapenlaan 136, NL-2512 HX The Hague, The Netherlands

Corresponding author: Ruben Huele; e-mail: huele@rulcml.leidenuniv.nl

Abstract

As data bases used in LCA are relatively large, it is generally not easy to form an opinion on its structure. A method of analysis for LCA data bases was developed to gain insight in this structure.

With the proposed analysis, a number of empty processes and emissions were found. Also a number of identical processes have been traced. It is concluded that the method of analysis forms a first step of a technique for finding errors in LCA data bases.

Keywords: Data base management, LCA; data bases, analysis; LCA data bases, correlation technique; LCI; Life Cycle Assessment, data bases; Life Cycle Inventory (LCI); spy plot, data base

1 Introduction/Objective

The reliability, and by implication the credibility, of Life Cycle Assessment (LCA) is largely dependent on the quality of the underlying data base. LCA is still under development and the theory and practice of LCA data base management is not yet as mature as for other fields.

The volume of the data involved in LCA studies makes it difficult to interpret all the information presented. The data and the results are usually summed, aggregated and normalised, to produce a small, comparable set of numbers. This aggregation, however, hides the information as present in the data base used. Moreover, the method of constructing the table inevitably introduces redundancy by structural multicollinearity (GLANTZ & SLINKER, 1990, p. 182).

Therefore, a graphic representation of data was sought first of all to visualise the structure of the data base and its internal linkages. Next, a simple factor analysis was carried out on both inventories and emissions/resources.

This paper presents a method for inspecting the calculated results of LCA data bases, with the help of correlation techniques. The large ETH 96 (FRISCHKNECHT, 1996) LCA data base is analysed with the help of Matlab.

* The tables appear in German as the original source is German.

2 Materials

The large ETH 96 (FRISCHKNECHT, 1996) LCA data base which results is used for the analysis. The data base contains data on processes related to the transformation and transport of energy. Each record describes the emissions, resource use and linkages to other processes. Conceptually, the data base consists of four tables, though the data are published in the form of 30 Excel worksheets. Two tables link names, categories, units and index numbers. The records in the third table describe the emissions, resources, and linkages per process, which form the raw data of the data base. The fourth table, named "ResEmis" is calculated from the third table and describes the total emissions/resources used for each inventory, including the data for the underlying processes it is linked with. In other words, the cradle to gate inventory table is calculated for each process. The table of results consists of 1182 records, each containing data on 616 emissions/resources used. This fourth table was used as material to analyse the internal structure of the data base.

3 Methods

3.1 Empty inventories and non-used emissions/resources

As a first step, empty rows were sought representing those inventories for which no emission and no resource data are provided. Next, empty columns were sought, representing substances not occurring in any of the available inventories. Empty columns and empty rows were deleted from the data structure before further calculations. In the remaining data structure, identical rows and columns were identified. It should be noted that empty cells were considered to contain zero.

3.2 Finding correlation coefficients

An analysis was made by correlating the inventories mutually, followed by such an analysis for emissions/resources. The product moment correlation coefficient, being a dimensionless covariance, is independent of the origins and

Int. J. LCA 3 (2) 114 – 118 (1998)

units of measurements (FELLER, 1968, p. 236) and can be calculated directly for the emissions/resources. For the analysis of the inventories, the data had to be normalised so as to reduce the effect of extreme outliers. This was done by substracting the mean and dividing by the standard deviation (JENNINGS & MCKEOWN, 1992, p. 181-186).

All calculations were carried out with Matlab, a commercial software package, that was originally developed to manipulate matrices.

4 Results

4.1 Empty inventories and non-used emissions/resources

The results of the filtering step are shown in **Figure 1** and in **Tables 1** and **2**. **Figure 1** represents the spy plot of the cells that contain zeros, either for all of the emissions/resources or for all of the inventories. The figure therefore shows the emissions/resources that do not appear in any inventory. These are presented in **Table 1**. Analogously, the figure shows those inventories that do not cause any emissions/resources. These are presented in **Table 2**.

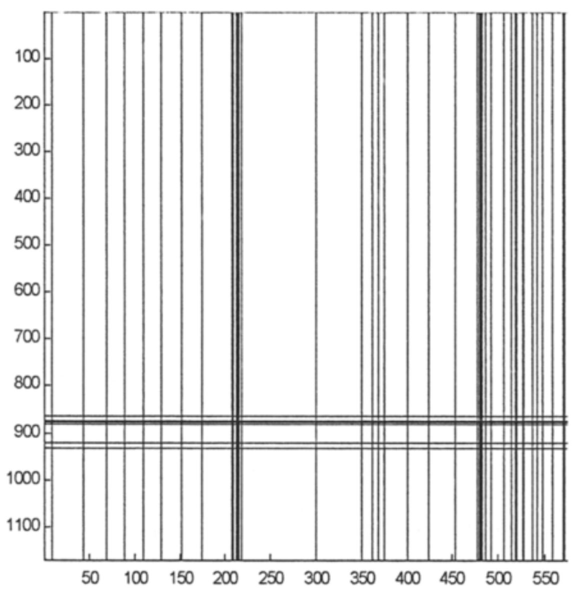

Fig. 1: Spy plot of the cells that contain no value, which are seen as zero. A vertical line represents an emission/resource that does not appear in any of the inventories, as shown in **Table 1**. A horizontal line represents an inventory without any emission/resource, as shown in **Table 2**

Almost all the emissions/resources mentioned in **Table 1** include an additional letter, like "p", "s", or "f". For a lot of these zero-emissions, the letter "p" is included which indicates that they do not apparently appear as process specific. Analogously, a lot of emissions are not emitted to seawater, indicated by the letter "s" or to freshwater, as indicated by the letter "f". The "Laerm" (noise) apparently has not been inventoried at all.

Table 1: The emissions/resources that do not appear in any inventory. Found as empty columns of the matrix

Flaeche Benthos III-IV
1,1,1-Trichlorethan p
BaP Benzo(a)pyren p
CH3Br p
Cycloalkane p
H 1211 Halon p
Laerm
Nitrate p
R113 FCKW p
R115 FCKW p
R141b H-FKW p
R142b H-FKW p
Radio. Aerosole p
Radio. Aktinide p
SF6 p
Barit f
Chlor. Dichlormonofluormethan f
Chlor. Methylenchlorid s
Chlor. Trichlormethan f
Glutaraldehyd in Wasser f
Ion Chrom-VI s
Ion Silizium s
Phosphor Verb. s
Rad. Aktinide f
....
Rad. U 238 s
Schwefel Gesamt f
Schwefelwasserstoff s
Tributylzinn TBT f
Sn in Boden

For the inventories mentioned in **Table 2** it is reported in the accompanying document that no emissions/resources appear. It has been documented that some provide useful byproducts so that the emissions/resources can be allocated to those byproducts that some processes are supposed to give no emissions/resources at all, and that no data are available for some processes.

Table 2: The inventories that contain no emissions/resources. Found as empty rows in the matrix

Deponierte Flugasche
Elektronikabfaelle
Infra Inertstoffdeponie
Ionentauscherharz in Sonderabfall
Katalysator in Sonderabfalldeponie
PV-Produktionsabfaelle in SAVA
PV/EVA-Zellenabfaelle
Schlamm Ionentauscher in Sonderabfall
Schweissstaub in Sonderabfallbehandlung

4.2 Finding correlation coefficients

The results from the correlation analysis are given in **Figures 2** and **3** and in **Tables 3** and **4**. The figures represent spy plots with mutual correlation factors of 1. Notice that the spy plot figures are symmetrical since the correlation of process x with y is equal to the one of y with x. Also notice that the diagonal should be filled entirely since the correlation with itself always equals 1 for any process. The diagonal,

however, has been left out in order to make the figure easier to read.

Figure 2 represents the spy plot of the correlation coefficients between emissions/resources where correlation coefficients prove to equal 1. A number of 23 emissions/resources prove to have a correlation coefficient of 1 with another emission/resource. Of these emissions/resources, 10 prove to differ over inventories. The correlation coefficient of 1 is then caused by a rounding of the numbers. In the other cases, the pairs turn out to be identical. These emissions are collected in **Table 3**. A number of these emissions appear in one or a limited number of the raw process data sheets in which they have the same value, either coincidentally or by assumption. Below, some of these are explained.

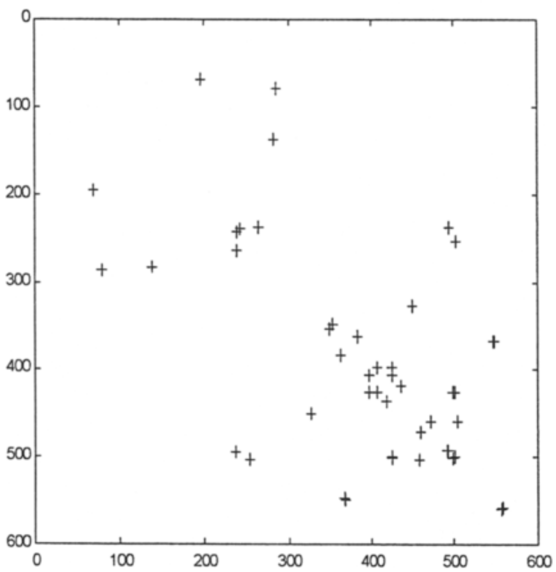

Fig. 2: Spy plot of the correlation coefficients between emissions/resources equal to 1, which is symmetrical in the diagonal axis. Those emissions/resources that appear identical for all inventories are shown in **Table 3**

Table 3: The emissions/resources that have the same value in every inventory. Found as identical lines in the matrix

Benzaldehyd s	=	Propionaldehyd s
Ca Calcium m	=	Si Silizium m
Radio. Pa234m p	=	Radio. Th234 p
Radio. Pb210 p	=	Radio. Po210 p
Chlor. HOCl s	=	Chlor. OCl s
Cyanide s	=	Ion Aluminium s
DOC f	=	Triethylen-Glykol f
DOC s	=	Triethylen-Glykol s
Ion Molybdaen s	=	Ion Vanadium s
Rad. Ba140 f	=	Rad. Sb122 f
Rad. C14 s	=	Rad. Cs134 s
Rad. Pb 210 f	=	Rad. Po 210 f
Rad. Ra 226 s	=	Rad. Ra 228 s

- The emissions of Calcium m and Silizium m appear only in one process, "Frachter Uebersee" (sea transport)
- The emissions DOC and Triethylen-Glykol appear only in the desulfurization of natural gas
- The emissions of Cyanide s and Aluminium s appear only in four refinery processes for car petrol

Figure 3 represents the spy plot of the correlation coefficients between inventories which equal 1. A number of 150 inventories prove to have a correlation coefficient of 1 with another inventory. Of these inventories, 10 prove to differ in 1 to 4 emissions/resources. The correlation coefficient of 1 is then caused by a rounding of the numbers. In the other cases, the pairs turn out to be identical. These inventories are presented in **Table 4**.

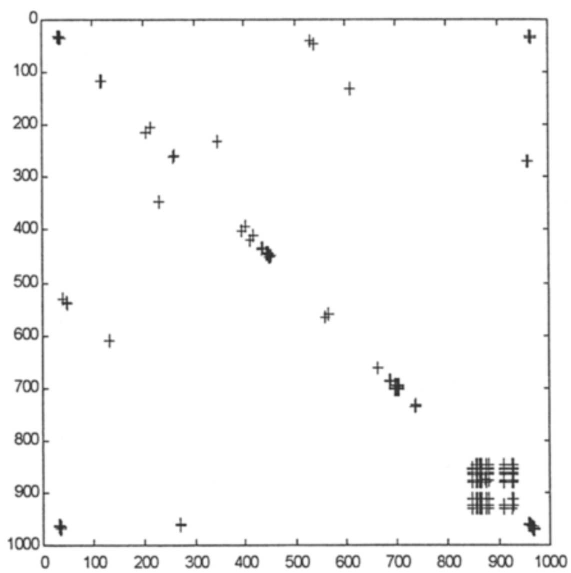

Fig. 3: Spy plot of the correlation coefficients between inventories equal to 1, which is symmetrical in the diagonal axis. Those inventories that appear identical for all emissions/resources are shown in **Table 4**

Inside a set of inventories, it can be seen that their subjects are quite similar. Presumably, some distinction should have been made between the raw process data, but this might not have been possible due to a lack of specific data. Below, some of the sets are explained from the raw process data.

For the "Strom ab Wasserkraft (Hydro power) processes, the report mentions the assumption that hydro power processes for Belgium, Spain, Greece, etc. are assumed to be equal to the average UCPTE. This average UCPTE process in turn is based on the average of hydro power in Austria, France, Italy, Germany and Switzerland, since these countries generate most of the hydroelectricity in Europe

Table 4: The inventories that cause identical amounts of emissions/resources. Found as identical rows in the matrix

Infra Strom Hochspannung - Bezug in CH Infra Strom Hochspannung - CH'93 Infra Strom Hochspannung - EWZ Infra Strom Hochspannung - NOK
Infra Strom Mittelspannung - Bezug in CH Infra Strom Mittelspannung - CH'93 Infra Strom Mittelspannung - EWZ Infra Strom Mittelspannung - NOK
Infra Strom Niederspannung - Bezug in CH Infra Strom Niederspannung - Bezug UCPTE
Stk fuer A Import Stk fuer A
Stk fuer I Import Stk fuer I
Stk fuer NL Import Stk fuer NL
Strom Niederspannung - Bezug in Ex-Ju Import Strom Niederspannung - Bezug in F
transp. europ. Stk frei UCPTE transportierte UCPTE Stk
Sauerstoff ab Luftzerlegung Stickstoff ab Luftzerlegung
Infra Hafenanlage Infra Ferntransport
Infra Baumaschine Infra Frontladerraupe
Infra UCPTE-Strom Ferntransport Infra Ferntransport CH'93 Infra Ferntransport EWZ Infra Ferntransport NOK
Strom oelthermisch B Strom oelthermisch L
Erdgas frei CH, GUS Erdgas frei UCPTE, GUS
Erdgas frei CH, N Erdgas frei UCPTE, N
Infra Foerdergas, Alg. Infra Foerdergas, D Infra Foerdergas, GUS
Infra produziertes Erdgas, Alg. Infra produziertes Erdgas, D Infra produziertes Erdgas, GUS Infra produziertes Erdgas, N Infra produziertes Erdgas, NL
Infra Transport Erdgas-Pipeline Infra Transport Erdgas-Pipeline Alg. Infra Transport Erdgas-Pipeline D Infra Transport Erdgas-Pipeline NL
Steinkohle aus Tagbau ab Bergwerk Stk ab Bergwerk Sued-Amerika
Uran anger.3.4% EURODIF Uran anger.3.4% fuer DWR F
Infra Speicherkraft UCPTE Infra Umwaelzwasserkraft UCPTE
Strom ab Wasserkraft B Strom ab Wasserkraft E Strom ab Wasserkraft Ex-Ju Strom ab Wasserkraft GR Strom ab Wasserkraft L Strom ab Wasserkraft NL Strom ab Wasserkraft P Strom ab Wasserkraft UCPTE
Infra Feuerung Fichte 300 Infra Feuerung Saegerei 300
Infra Feuerung Fichte 50 Infra Feuerung Saegerei 50
Abfaelle in Inertstoffdeponie Bausperrgut in Inertstoffdeponie Beton in Inertstoffdeponie Deckfarbe in Inertstoffdeponie Erdgasleitungen in Inertstoffdeponie Glas in Inertstoffdeponie Kalksteinrueckstaende in Inertstoffdeponie Kupfer in Inertstoffdeponie Mineralwolle in Inertstoffdeponie Stahl in Inertstoffdeponie Zeolithe in Inertstoffdeponie
Infra Ferntransport CH'93 Infra Ferntransport EWZ Infra Ferntransport NOK
Infra Strom Hochspannung - CH'93 Infra Strom Hochspannung - EWZ Infra Strom Hochspannung - NOK
Infra Strom Mittelspannung - CH'93 Infra Strom Mittelspannung - EWZ Infra Strom Mittelspannung - NOK
Infra Strom Niederspannung - CH'93 Infra Strom Niederspannung - EWZ Infra Strom Niederspannung - NOK

- For the "Stickstoff, Sauerstoff ab Lufterzeugung" (air separation processes), the report mentions the identical allocation of these two processes
- For the "Inertstoffdeponie" (inert waste dump) the report mentions that the different processes are assumed to be equal

5 Conclusion

5.1 Filtering empty inventories and non-used emissions/resources

The use of spy plots forms a useful way to visualise the presence of empty inventories and non used emissions/resources in an LCA data base.

5.2 Finding correlation coefficients

The use of correlation coefficients assists in finding identical inventories and identical emissions/resources. Correspondences can be explained from the raw data. Some of the correspondences are due to explicit assumptions, such as the equal allocation for two products. Some of the correspondences are due to a lack of specific raw data for resembling processes. Some of the correspondences are due to the fact that some emissions appear at only one specific process in the raw data base.

5.3 General

The analysis has pointed out some peculiarities of the raw data base considered. Most of these peculiarities are explicitly mentioned and explained in the report.

6 Discussion and Future Outlook

1. High correlations in the data base indicate a redundancy of information. This redundancy in the "Resemis" table is predominantly caused by the way the table is constructed. The examination of possible correlations in the raw data in the constituting tables is planned for future work.
2. In this case, the causes of the redundancy can be found in the elaborate documentation that is available for the raw data base. Measuring the redundancy of the Resemis table mainly has a heuristic value. For the evaluation of other data bases for which documentation is lacking or incomplete, estimating the redundancy will be essential.
3. The problem of evaluating the information content of a data base is equivalent to multivariate factor analysis. The processes are interpreted as the independent variables and the environmental data as the dependent variables. Future analysis of LCA data bases might well benefit from the techniques developed in factor analysis.
4. Finding the redundancies, it might be possible to identify the dominant processes containing most information and, by implication, the processes most suited for a quick-and-dirty LCA. In the terms of factor analysis, finding the dominant processes is equivalent to finding the principal components, that is the eigenvectors with the highest eigenvalue.

Acknowledgement

We thank the anonymous referent of the first draft for the comments made. Nathalie van der Linden stimulated us to pay attention to factor analysis.

8 References

Feller, W.: An Introduction to Probability Theory and its applications. Third Edition, Volume 1. John Wiley & Sons, New York, 1968

Frischknecht, R.; Hofstetter, P.; Knoepfel, I.; Dones, R.; Zollinger, E.: Ökoinventare von Energiesysteme, Grundlagen für den ökologischen Vergleich von Energiesystemen und den Einbezug von Energiesystemen in Ökobilanzen für die Schweiz. 3. Auflage. ESU-ETH Zürich, GS-PSI Villingen, 1996. In German

Glantz, Stanton A.; Slinker, B.K.: Primer of Applied Regression and Analysis of Variance. McGraw-Hill Inc, New York, 1990

Hair, J.F.Jr.; Anderson, R.E.; Tatham, R.L.: Multivariate Data Analysis. With readings. Second Edition. McMillan Publishing Company, New York, 1987

Jennings, A.; McKeown, J.J.: Matrix Computation. Second Edition. John Wiley & Sons, New York, 1992

Matlab: High-Performance Numeric Computation and Visualization Software, Reference Guide. The Math Works Inc, Natick, Mass., 1996

Conference Announcements

The Second International Conference on **LCA in Agriculture, Agro-Industry and Forestry** will be held at the Hotel Palace in **Brussels, Belgium, on December 3-4, 1998.** The first day will focus on the potential uses of Life Cycle Assessment (LCA) in agriculture, agro-industry, biotechnology and forestry. Methodological problems which arise when LCA is applied in these areas will be discussed. The second day will focus on case studies and the use of LCA in decision making processes.

The meeting is open, and abstracts for platform and poster presentations are invited. Abstracts should be written in English and not exceed 300 words. They should be sent to the address below.

More information can be obtained from:
Mieke Engelen
VITO, Product and Process Assessment
Boeretang 200
B-2400 Mol, BELGIUM
Phone: +32 14 33 58 60
Fax: +32 14 32 11 85
E-mail: engelenm@vito.be

First Announcement

9th Annual Meeting of SETAC-Europe:

Quality of Life and Environment in Cultured Landscapes

University of Leipzig, Germany, May 25-29, 1999

Main Themes
Ecosystem Monitoring and Analysis
Environmental Hazard and Risk
Normative and Regulatory Strategies
Ecological Economy and Environmental Technology

The technical programme includes invited lectures addressing general aspects of the above-mentioned main themes, platform and poster sessions, short courses to provide educational opportunities, and special symposia discussing selected items like environmental consequences of the use of antibiotics, and spatial aspects of exposure and risk assessment. In addition, an exhibition area will be available for presentations from industry, consultants, manufacturers and publishers.

Special emphasis will be given to translating scientific methods and results into practical tools that are needed for a sustainable development of societies on a regional and global scale. Both experimental and theoretical work is welcome for presentation and discussion, and the wide range of oral and poster sessions will address recent as well as traditional topics in the environmental sciences.